caBIG Workspace
Vocabularies and Common Data Elements

caDSR

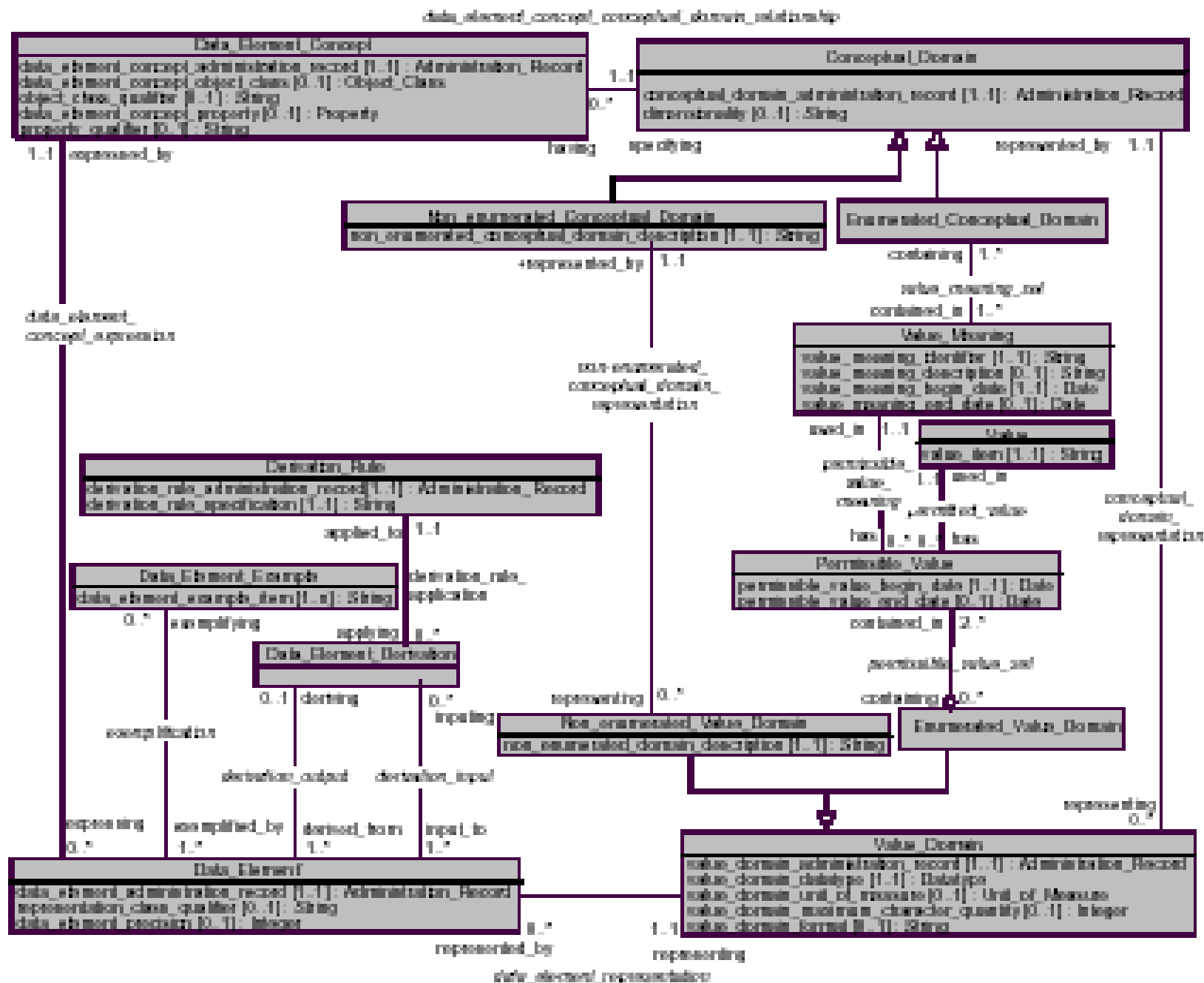March 30, 2004
Denise Warzel
NCI, Center for Bioinformatics

CDE and Vocabulary Workspace Definition

▸ Support for Domain Workspaces in development and use of Common Data Elements (CDE) and vocabulary.

▸ Define needs for the development of software systems/tools to support development and implementation of CDEs and vocabulary.

▸ *Support for the development of CDEs* and vocabulary content for use in caBIG. Encompass activities with a common focus, addressing identified Cancer Center needs
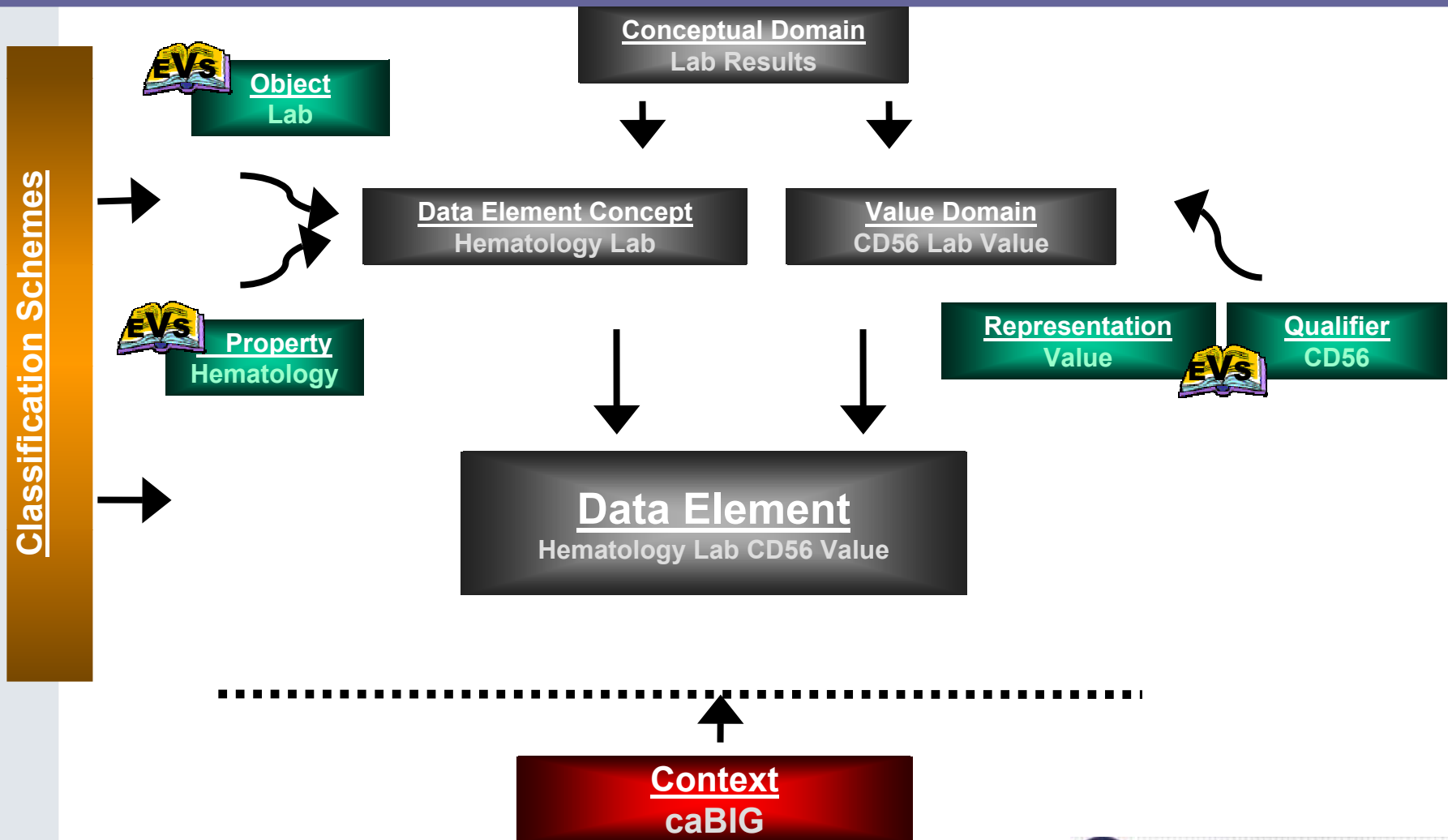
Agenda

▶ **Introductions**

▶ **Overview ISO 11179 metamodel**

▶ **Demo Browser**
→ **Navigation**
→ **Queries**
→ **Export**

▶ **Demo Admin Tool**

→ **Create Administered Components**

▶ **Demo Curation Tool**
→ **Create DEC, VD, DE**

ISO Consolidated Metamodel

ISO/IEC 11179 Model

Review caDSR and CDEs

caDSR:

➢Based on the ISO 11179 standard for metadata registries

➢Designed to integrate caCORE infrastructure

➢Supports the development of Data Elements that are used as metadata descriptors for NCI-sponsored research

Review caDSR and CDEs

- Metadata is managed by Contexts

- ISO 11179 Administered Components include:
  - Data Elements (data element concept + value domain)
  - Data Element Concepts (comprised of object class + property)
  - Value Domains (representation class)
  - Conceptual Domains (all possible values)
  - Classification Schemes (organization structure for groups of components)
    - Types of classification schemes include Key Words, Thesuari,Taxonomies, Ontologies

Review caDSR and CDEs

Tools to _Develop_, _Access_ and _Consume_

Content Development:

- **Administration Tool**
  - Curate all types of ISO administered components
- **Curation Tool**
  - Curate DEC, DE, VD, and indirectly, OC, P and Rep terms
- **Excel Loader**
- **UML Loader**

Access:

- **CDE Browser**

Consumption:

- **CDE Browser Download** (Excel, XML)
- **Case Report Form Review**
  - Case Report Form (CRF) Loader (Excel)
  - Compliance Review Tool (CRT) (Excel)
  - CDE Compliance Review Response Tool (CCRR)

- **caBIO APIs**
  - HTTP-XML, SOAP-XML and Java

Demos

Access

CDE Browser

→ Public except for Form Builder (2.1): userid and password

Curation

Administration Tool

→ Userid and password
→ UML Model and Excel Batch Loaders

CDE Curation Tool

→ Userid and password

Send Request for caDSR Account to: ncicb@pop.nci.nih.gov

Or http://ncicb.nci.nih.gov/core/caDSR and follow the link to NCICB Application Support

caBIG   *cancer Biomedical Informatics Grid*

Import

– Excel Loader

- Data transformed from Excel spreadsheet into caDSR metadata

- Spreadsheet template available by request from NCICB help desk

– UML Loader

- UI accesses from Admin Tool

- Data transformed from UML Model into caDSR metadata

- Multi-step process

- User Interface allows users to set metadata defaults for batch load

Recommended Readings

- ▶ ISO 11179 Part-1, Framework and specifications, Section Appendix A.2, Classifiations

- ▶ ISO 11179 Part-2, Classification of data elements

- ▶ ISO 11179 Part-3, Basic attributes of standardized data elements

- ▶ ISO 11179 Part-4, Rules and guidelines for formation of definitions

- ▶ ISO 11179 Part-5, Naming and identification principals for data elements

- ▶ ISO 11179 Part-6, Registration of data elements

- ▶ ISO 20944-02 Metadata registry interoperability and bindings – Part 2: Attribute mapping, Section 4, Identifier mangling conventions

- ▶ ISO 20943- Part 1: Metadata registry content consistency – Part 1: Data Elements. Section 6: Bottoms-up approach Section 7: Top-down approach

- ▶ ISO 20943- Part 3: Metadata registry content consistency – Part 1: Value Domains

Operating Environments

▸ Database Repository
  – Oracle 8i

▸ Administration Tool
  – Oracle PL/SQL, Oracle 9i Application Server

▸ CDE Browser
  – Java, Oracle 9i Application Server

▸ CDE Curation Tool
  – Jakarta Tomcat

▸ Case Report Form (CRF) Loader
  – Visual Basic, MS Excel GUI, Oracle 8i Client

▸ Compliance Review Tool (CRT)
  – Visual Basic, MS Excel GUI, Oracle 8i Client

▸ CDE Compliance Review Response Tool (CCRR)
  – Java, JSP, Java Beans, Oracle 9i Application Server

caDSR Team

- ▸ NCI
  - – Peter Covitz
  - – Denise Warzel
  - – caDSR Context Administrators

- ▸ ScenPro
  - – Bill McCurry
  - – Mike Gately
  - – Ken Sims
  - – Mark Swenholt
  - – Jennifer Brush
  - – Tom Phillips
  - – Mike Holck
  - – Smita Hastak
  - – Linda Braswell

- ▸ Oracle
  - – Edmond Mulaire
  - – Ram Chilukuri
  - – Jasur Ismukhamedov
  - – Prerna Aggarwal
  - – Hyun Kim
  - – Wendy Verhoef

- ▸ Science Applications International Corporation (SAIC)
  - – Kathleen Gundry
  - – Tommie Curtis
  - – Brenda Maeske
  - – Mead Walker

Questions?

- caDSR Home Page
    - http://ncicb.nci.nih.gov/core/caDSR

- Administration Tool
    - http://cadsr-prod.nci.nih.gov

- CDE Curation Tool
    - http://ncicb.nci.nih.gov/cdecurate

- CDE Browser

- http://ncicb.nci.nih.gov/cdebrowse

- CCRR:
    - http://ccrr-prod.nci.nih.gov/ccrr

- WarzelD@mail.nih.gov

- 303-722-9446